HoGent

BEDRIJF
EN
ORGANISATIE

POD II - Deel I: Kansrekening

Jens Buysse (Jens.Buysse@hogent.be)
Harm De Weirdt (Harm.DeWeirdt@hogent.be)
Wim Goedertier (Wim.Goedertier@hogent.be)
Stijn Lievens (Stijn.Lievens@hogent.be)
Lieven Smits (Lieven.Smits@hogent.be)

Deel I: Kansrekening

Inhoud

Gebeurtenissen/toevalsverschijnselen

- De KANSREKENING houdt zich bezig met de studie van *gebeurtenissen* of *toevalsverschijnselen*.
- Bij een gebeurtenis is de individuele uitkomst onzeker.
  Bv. bij het gooien met een dobbelsteen weet je niet hoeveel ogen de volgende worp zal tonen.
- Bij een groot aantal herhalingen ontstaat een regelmatige verdeling van relatieve frequenties.
  Bv. bij een "eerlijke" dobbelsteen verwacht je dat, na een groot aantal worpen, elk van de zes mogelijkheden ongeveer even vaak voorkwam.
- De kansrekening heeft toepassingen in een groot aantal zeer uiteenlopende gebieden.

Voorbeeld

Twee gokspelen van Chevalier de Méré.

1. Spel 1: inzetten op minstens één zes bij 4 keer gooien met een dobbelsteen.
   Hierbij won Chevalier de Méré.
2. Spel 2 (uitdagender): dubbelzes te gooien bij een worp met twee dobbelstenen. Inzetten op minstens één dubbelzes bij $4 \times 6 = 24$ keer gooien.
   Redenering: kans op dubbelzes is $1/6$ van kans op enkele zes (bij één worp). *E*rgo 6 keer meer worpen nodig.

Helaas voor hem was/is deze redenering **fout**, en verloor hij zwaar.

Het begrip "kans"

- Uitspraak: "Als ik een evenwichtig muntstuk tos, dan is de kans op munt 0,50."
- Hoe kan deze uitspraak geverifieerd worden?
- Antwoord: aan de hand van een groot aantal experimenten!
- Het begrip "kans" komt overeen met de relatieve frequentie *op lange termijn*.
- De beschrijving van het opgooien van een muntstuk omvat twee onderdelen:
    1. het opstellen van een lijst van mogelijke uitkomsten;
    2. het bepalen van de kans van elke uitkomst.

Universum of uitkomstenruimte

Definitie

Het UNIVERSUM *of de* UITKOMSTENRUIMTE *van een experiment is de verzameling van alle mogelijke uitkomsten van dit experiment en wordt genoteerd met* Ω.

Opmerkingen

- De uitkomstenruimte moet *volledig* zijn: elke mogelijke uitkomst van een experiment moet tot Ω behoren.
- Bovendien moet elke uitkomst van een experiment overeenkomen met *juist één* element van Ω.
- Samengevat: na het uitvoeren van een experiment is het mogelijk om eenduidig aan te geven welk element van Ω zich heeft voorgedaan.

Voorbeelden

- bij het opgooien van een muntstuk is $\Omega = \{\text{Kop}, \text{Munt}\}$;
- bij het opgooien van 1 dobbelsteen is $\Omega = \{1, 2, 3, 4, 5, 6\}$;
- bij het opgooien van 2 dobbelstenen wordt dit

  $$\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\} = \{(1,1),(1,2),\ldots,(6,6)\};$$

- wanneer men willekeurig een punt kiest in de eenheidscirkel dan is

  $$\Omega = \{(x,y) \in \mathbb{R}^2 \,|\, x^2 + y^2 \leq 1\};$$

- bij het trekken van één kaart uit een (klassiek) kaartspel is

  $$\Omega = \{\diamondsuit 1, \ldots, \diamondsuit 13, \clubsuit 1, \ldots, \clubsuit 13, \heartsuit 1, \ldots, \heartsuit 13, \spadesuit 1, \ldots, \spadesuit 13\}.$$

Gebeurtenis

- Ω is een opsomming van *alle* mogelijke uitkomsten van een experiment.
- Bv. bij viermaal gooien van een dobbelsteen is $\Omega = \{1, 2, 3, 4, 5, 6\}^4$, maar Chevalier de Méré was enkel geïnteresseerd of er minstens één zes was gegooid.
  i.e. in de gebeurtenis: "er werd minstens één zes gegooid"
- Deze gebeurtenis kan geschreven worden als een *deelverzameling* van Ω.

Gebeurtenis: voorbeeld

Bij het éénmaal opgooien van een dobbelsteen is $\Omega = \{1, 2, 3, 4, 5, 6\}$. We kunnen spreken over de volgende gebeurtenissen:

- de uitkomst is het getal 1, of $A_1 = \{1\}$;
- de uitkomst is een even getal, of $A_2 = \{2, 4, 6\}$;
- de uitkomst is even en is niet groter dan 3, of $A_3 = \{2, 4, 6\} \cap \overline{\{4, 5, 6\}}$;
- de uitkomst is niet even, of $A_4 = \overline{\{2, 4, 6\}}$.

Gebeurtenis: definitie

Definitie

Een GEBEURTENIS *is een deelverzameling van de uitkomstenruimte. Een* ENKELVOUDIGE *of* ELEMENTAIRE *gebeurtenis is een singleton; een* SAMENGESTELDE GEBEURTENIS *heeft cardinaliteit groter dan 1.*

Gebeurtenissen die geen gemeenschappelijke uitkomsten hebben noemt men DISJUNCT.

Disjuncte gebeurtenissen kunnen dus nooit samen voorkomen.

Wanneer de gebeurtenissen A en B disjunct zijn dan geldt $A \cap B = \emptyset$.

Voorbeelden

Bij het gooien van een dobbelsteen geldt:

1. de gebeurtenis $A_1 = \{1\}$ is enkelvoudig;
2. de gebeurtenis $A_2 = \{2, 4, 6\}$ is samengesteld;
3. de gebeurtenissen A_1 en A_2 zijn disjunct;
4. de gebeurtenissen A_2 en A_4 zijn disjunct.

Nieuwe gebeurtenissen van bestaande

Startend met de gebeurtenissen A en B kan men de volgende gebeurtenissen vormen:

- A **of** B, of wiskundig genoteerd $A \cup B$;
- A **en** B, of wiskundig genoteerd $A \cap B$;
- **niet** A, of wiskundig genoteerd $\overline{A}$.

Opmerkingen

- Door inductie leidt men gemakkelijk af dat de unie van n (met $n \in \mathbb{N}$) gebeurtenissen A_1 t.e.m. A_n eveneens een gebeurtenis is.
- Idem voor de doorsnede van gebeurtenissen.
- Voor sommige toepassingen is het nodig om ook (aftelbaar) oneindige unies en doorsnedes te beschouwen.

Oefeningen

1. Een dobbelsteen wordt éénmaal opgegooid. Geef de volgende gebeurtenissen door opsomming:
   1.1 er werd een even getal gegooid;
   1.2 er werd een priemgetal gegooid;
   1.3 er werd een even getal of een priemgetal gegooid;
   1.4 er werd een even getal en een priemgetal gegooid.
2. Er wordt tweemaal gegooid met een dobbelsteen. Omschrijf de volgende gebeurtenissen:
   2.1 er wordt 1 gegooid bij de eerste worp;
   2.2 de twee worpen vertonen hetzelfde aantal ogen;
   2.3 het minimum van de worpen is drie.
3. Geef de volgende gebeurtenissen voor het experiment in Voorbeeld 1.11:
   3.1 er zijn juist vijf worpen nodig;
   3.2 er zijn hoogstens vijf worpen nodig;
   3.3 er zijn minstens vijf worpen nodig.

Kansen en kansruimte: Inleiding

Doel: koppel aan elke gebeurtenis A een getal dat uitdrukt hoe waarschijnlijk het is dat A voorkomt bij het uitvoeren van een experiment. We noemen dit getal de KANS of WAARSCHIJNLIJKHEID van A, en we noteren deze kans als $\mathbb{P}(A)$.

Om ervoor te zorgen dat de theorie realistisch is moet de functie $\mathbb{P}$ aan bepaalde voorwaarden voldoen (zie volgende slide).

Kansruimte

Definitie

Het toekennen van kansen aan gebeurtenissen dient aan de volgende drie regels te voldoen.

1. *Kansen zijn steeds positief:* $\mathbb{P}(A) \geq 0$ *voor elke* A.
2. *De uitkomstenruimte heeft kans 1:* $\mathbb{P}(\Omega) = 1$.
3. *Wanneer* A *en* B *disjuncte gebeurtenissen zijn dan is*

   $$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

   Dit noemt men de SOMREGEL.

Wanneer de functie $\mathbb{P}$ *aan de bovenstaande eigenschappen (axioma's) voldoet dan noemt men het drietal* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *een* KANSRUIMTE.

Eigenschappen van kansen

Stelling

Kansen voldoen aan de volgende eigenschappen:

1. *Voor elke gebeurtenis A geldt dat $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.*
2. *De onmogelijke gebeurtenis heeft kans nul: $\mathbb{P}(\emptyset) = 0$.*
3. *Als $A \subseteq B$, dan is $\mathbb{P}(A) \leq \mathbb{P}(B)$; i.h.b. geldt $\mathbb{P}(A) = \mathbb{P}(B) - \mathbb{P}(B \setminus A)$.*
   Dit is een soort van verschilregel.
4. *De* UITGEBREIDE SOMREGEL *is:*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Eindig universum

- Als Ω eindig is, i.e. $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_n\}$, dan volgt uit de axioma's van kansruimte dat

$$\sum_{i=1}^{n} \mathbb{P}(\omega_i) = 1.$$

- Waneer de elementaire gebeurtenissen allen *even waarschijnlijk* zijn, i.e. $\mathbb{P}(\omega_i) = 1/n$, dan bekomt men de FORMULE VAN LAPLACE:

$$\mathbb{P}(A) = \frac{\#(A)}{\#(\Omega)}.$$

Eindig universum: voorbeeld

Men trekt *willekeurig* een kaart uit een klassiek kaartspel. Wat is de kans dat men een aas trekt?
De gebeurtenis "men trekt een aas" wordt voorgesteld door de volgende verzameling

$$A = \{\diamondsuit 1, \spadesuit 1, \heartsuit 1, \clubsuit 1\}.$$

De kardinaliteit van het universum Ω is uiteraard 52. Volgens de formule van Laplace krijgen we dus

$$\mathbb{P}(A) = \frac{\#(A)}{\#(\Omega)} = \frac{4}{52} = \frac{1}{13}.$$

Oefeningen

1. Men gooit éénmaal met een eerlijke dobbelsteen.
   1.1 Wat is de kans op een even getal?
   1.2 Wat is de kans op een oneven getal?
   1.3 Wat is de kans op een priemgetal?
2. Wat is de kans om met twee dobbelstenen meer dan acht te gooien?
3. Een duistere urne bevat 6 blauwe en 4 rode bollen die verder identiek zijn. Men trekt één na één, twee bollen uit de urne met teruglegging van de getrokken bol. Wat is de kans dat men twee blauwe bollen heeft getrokken?
4. Wat is de kans om bij de Belgische lotto zes juiste cijfers te hebben? Er zitten 45 ballen in de trommel en er worden er zes getrokken zonder teruglegging.

Oefeningen

5. Als huishoudens groepen mensen zijn die samenleven, dan vindt men voor de kansen op het aantal leden van een huishouden respectievelijk:

Omvang	1	2	3	4	5	6	7 of meer
Kans	$0,236$	$0,320$	$0,181$	$0,156$	$0,069$	$0,024$	$0,014$

   Het kleine aantal huishoudens met meer dan 7 leden werd samengevoegd in de laatste categorie. Wat is de kans dat er meer dan twee mensen in een huishouden zijn?

6. Een toegangscode tot een (amateurs) computersysteem bestaat uit drie letters die door een randomgenerator worden toegewezen.
   6.1 Wat is de kans dat een code wordt toegewezen zonder de letter x?
   6.2 Wat is de kans dat een code wordt toegewezen waarvan alle letters verschillend zijn?

Oefeningen

7. Bij een vervalst muntstuk is $\mathbb{P}(\mathsf{K}) = 6/10$. Wat is dan de kans op “munt”?
8. Veronderstel dat

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$$

met

$$\mathbb{P}(\omega_1) = \frac{3}{10},\ \mathbb{P}(\omega_2) = \frac{2}{10},\ \mathbb{P}(\omega_3) = \frac{1}{10},\ \mathbb{P}(\omega_4) = \frac{1}{10} \text{ en } \mathbb{P}(\omega_5) = \frac{3}{10}.$$

Wat is $\mathbb{P}(\{\omega_2, \omega_4, \omega_5\})$?

9. Twee zijden van een dobbelsteen dragen het cijfer 1, drie dragen het cijfer 2 en één zijde het cijfer 3. Wat is de kans op een oneven getal als men aanneemt dat elke zijde even waarschijnlijk is?

Oefeningen

10. Men vervalst een dobbelsteen zodanig dat de kans op een bepaald cijfer evenredig is met dit cijfer.
    - 10.1 Geef de elementaire kansen.
    - 10.2 Wat is de kans op een even getal?
    - 10.3 Wat is de kans op een oneven getal?
    - 10.4 Wat is de kans op een priemgetal?
    - 10.5 Wat is de kans op een even getal of een priemgetal?
    - 10.6 Wat is de kans op een oneven priemgetal?
11. Een doos bevat 10 lampen. Bij vergissing worden 4 afgekeurde lampen mee verpakt. Men neemt lukraak drie lampen uit de doos. Wat is $\mathbb{P}(\omega_i)$ waarbij ω_i de elementaire gebeurtenis voorstelt dat exact i $(0 \leq i \leq 3)$ van de drie gekozen lampen defect zijn?

Oefeningen

12. Hoe luidt de algemene uitgebreide somregel voor drie gebeurtenissen? Geef m.a.w. een uitdrukking voor

$$\mathbb{P}(A_1 \cup A_2 \cup A_3).$$

    Kun je de formule veralgemenen tot n gebeurtenissen?

13. Aan de Hogeschool Gent mogen (moeten?) 500 studenten 3 keuzevakken kiezen. Van de 500 studenten volgen er 320 cursus A, 190 volgen cursus B en 280 volgen cursus C. Er zijn 80 studenten die A en B volgen, 200 die A en C volgen terwijl 60 studenten B en C volgen.
    Bereken de kans dat een willekeurig gekozen student
    13.1 de drie cursussen volgt;
    13.2 A volgt maar niet B;
    13.3 C volgt maar niet B;
    13.4 A of C volgt maar niet B;
    13.5 A volgt maar B noch C.

Voorwaardelijke kansen

Stel A is de gebeurtenis “er is een inbraak aan de gang”

Stel B is de gebeurtenis “het alarm gaat af”.

Het lijkt nu logisch dat het voorkomen van B een *invloed* heeft op de waarschijnlijkheid dat A voorkomt. (M.a.w. als je weet dat B voorkomt, dan ga je wellicht de waarschijnlijkheid dat A voorkomt bijstellen.)

De kans dat A voorkomt als gegeven is dat B voorkomt wordt genoteerd als $\mathbb{P}(A \mid B)$.

Voorwaardelijke kansen: formule

Om een formule te vinden voor $\mathbb{P}(A \mid B)$ denken we frequentistisch.

Wanneer we een experiment herhaaldelijk uitvoeren dan kunnen we telkens het al dan niet voorkomen van A en/of B observeren.

Veronderstel nu dat we enkel geïnteresseerd zijn in dié experimenten waarvoor B voorkomt. De verhouding van het aantal keer dat A voorkomt voor deze experimenten is $N(A \cap B)/N(B)$, aangezien B voor deze experimenten steeds voorkomt. Er geldt nu echter dat

$$\frac{N(A \cap B)}{N(B)} = \frac{N(A \cap B)/N}{N(B)/N}.$$

Als we deze verhoudingen nu als een waarschijnlijkheid beschouwen dat zien we dat het getal $\mathbb{P}(A \mid B)$ redelijkerwijs moet gedefinieerd worden als $\mathbb{P}(A \cap B)/\,\mathbb{P}(B)$.

Voorwaardelijke kans: definitie

Definitie

Als $\mathbb{P}(B) > 0$*, dan is de* VOORWAARDELIJKE KANS *dat* A *voorkomt als gegeven is dat* B *voorkomt gedefinieerd als:*

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Opmerking

We spreken $\mathbb{P}(A \mid B)$ uit als "de kans op A gegeven B".

Voorbeeld

Er wordt tweemaal gegooid met een eerlijke dobbelsteen. Wat is de waarschijnlijkheid dat de som van beide ogen 7 of meer is als gegeven is dat de eerste worp een twee opleverde?

Intuïtitief is het duidelijk dat de gevraagde kans gelijk is aan $2/6$ aangezien we bij de tweede worp 5 of 6 moeten gooien.

Bekom hetzelfde resultaat door de definitie van voorwaardelijke kans toe te passen.

Onafhankelijke gebeurtenissen

Definitie

Twee gebeurtenissen A en B zijn ONAFHANKELIJK *als*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B).$$

Opmerking

Als A en B onafhankelijk zijn van elkaar, dan heeft het voorkomen van de ene gebeurtenis geen invloed op het al dan niet voorkomen van de andere gebeurtenis.

Onafhankelijke vs. disjuncte gebeurtenissen

Opmerking: Verwar *onafhankelijke* gebeurtenissen niet met *disjuncte* gebeurtenissen. Twee disjuncte gebeurtenissen A en B met $\mathbb{P}(A) > 0$ en $\mathbb{P}(B) > 0$ zijn steeds *afhankelijk*, want

$$\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0 \neq \mathbb{P}(A)\,\mathbb{P}(B).$$

Dit is ook logisch; als je bv. weet dat A voorkomt dan weet je zeker dat B niet voorkomt.
De kennis over het al of niet voorkomen van A heeft dus zeker een invloed op je kennis van het al dan niet voorkomen van B.

Voorbeeld

Beschouw een standaard kaartspel waaruit willekeurig een kaart wordt getrokken. We stellen door A de gebeurtenis voor dat de getrokken kaart een $\diamondsuit$ is, terwijl B de gebeurtenis voorstelt dat de getrokken kaart een aas is. We gaan nu na of A en B onafhankelijk zijn.

Intuïtief lijkt dit zo te moeten zijn omdat het feit dat je weet dat de kaart een $\diamondsuit$ is geen invloed heeft op het al dan niet een aas zijn aangezien alle 4 de kaarttypes juist één aas hebben.

Ga dit formeel na.

Voorbeeld afhankelijke gebeurtenissen

Ga formeel na dat de gebeurtenissen A en B uit Voorbeeld 1.20 afhankelijk zijn.
We hebben gebeurtenis A: de som is 7 of meer.
En gebeurtenis B: er werd twee gegooid bij de eerste worp.

Voorbeeld

Beschouw een familie met 3 kinderen waarbij elk kind dezelfde kans heeft om een jongen of een meisje te zijn, onafhankelijk van het geslacht van de andere kinderen. Beschouw de volgende gebeurtenissen:

$$A = \{\text{alle kinderen hebben hetzelfde geslacht}\},$$
$$B = \{\text{er is hoogstens één jongen}\},$$
$$C = \{\text{er is minstens één jongen en minstens één meisje}\}.$$

Welke koppels gebeurtenissen zijn onafhankelijk?

Kettingregel voor gebeurtenissen

Theorem (Kettingregel)

Wanneer A_1 *t.e.m.* A_n n *gebeurtenissen zijn waarvoor*

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$$

dan geldt

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\,\mathbb{P}(A_2 \mid A_1)\,\mathbb{P}(A_3 \mid A_1 \cap A_2) \cdots \mathbb{P}(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}$$

De waarschijnlijkheid van de doorsnede kan dus geschreven worden als een product van voorwaardelijke waarschijnlijkheden. Dit noemt men ook de KETTINGREGEL *of* PRODUCTREGEL*.*

Toepassing: de verjaardagsparadox

Gebruik de kettingregel om een formule te vinden voor de kans dat k willekeurig gekozen personen allemaal een verschillende verjaardag hebben.
Toon in het bijzonder aan dat deze kans gegeven wordt door:

$$\prod_{j=1}^{k} \frac{366-j}{365}.$$

Voor welke waarde van k is deze kans voor het eerst kleiner dan 50%?

Oefeningen

1. Een vaas bevat 20 witte en 30 zwarte bollen. Men trekt willekeurig drie bollen uit de vaas. Bereken de kans om 3 witte bollen te trekken
   1.1 met teruglegging;
   1.2 zonder teruglegging.
2. De kans dat van een koppel de man nog leeft over 10 jaar is $1/4$. De kans dat de vrouw dan nog leeft is $1/3$. (De gegevens komen uit zogenaamde *sterftetafels*, een veel gebruikt instrument in het gebied van de (levens)verzekeringen.) Bereken de kans dat
   2.1 beiden nog leven over 10 jaar;
   2.2 ten minste één van de twee nog leeft over 10 jaar;
   2.3 geen van beiden nog leeft;
   2.4 alleen de vrouw nog leeft.

   Welke veronderstelling heb je gemaakt tijdens je berekeningen?

Oefeningen

3. Doos A bevat 8 onderdelen, waarvan er 3 defect zijn. Doos B bevat 5 onderdelen, waarvan er 2 stuk zijn. Men neemt lukraak een onderdeel uit elke doos.
   3.1 Wat is de kans dat beide onderdelen intact zijn?
   3.2 Wat is de kans dat juist één van de twee onderdelen defect is?
   3.3 Als juist één van de twee onderdelen defect is, wat is dan de kans dat het defecte stuk uit doos A komt?
4. De kansen dat 3 boogschutters een bepaald doelwit raken zijn respectievelijk $1/6, 1/4$ en $1/3$. Iedere boogschutter schiet één keer naar het doel.
   4.1 Wat is de kans dat juist 1 van de 3 het doel raakt?
   4.2 Als juist 1 schutter het doel raakte, wat is dan de kans dat het de eerste schutter was?

5. Ga het volgende formeel na: als A en B onafhankelijke gebeurtenissen zijn, dan zijn ook $\overline{A}$ en B onafhankelijke gebeurtenissen.
   Hieruit volgt dan ook onmiddellijk dat A en $\overline{B}$ onafhankelijk zijn als A en B onafhankelijk zijn. Hetzelfde geldt dan uiteraard voor $\overline{A}$ en $\overline{B}$.

6. Veronderstel dat A en B gebeurtenissen zijn waarvoor $\mathbb{P}(A) = 1/4$, $\mathbb{P}(A \cap B) = 1/6$ en $\mathbb{P}(B) = p$.
    6.1 Bereken p als men ervan uitgaat dat A en B onafhankelijk zijn.
    6.2 Bereken p als $B \subset A$.
7. Veronderstel dat A en B onafhankelijke gebeurtenissen zijn met $\mathbb{P}(A) = 1/2$ en $\mathbb{P}(A \cap B) = 1/3$. Bereken dan
    7.1 $\mathbb{P}(B)$;
    7.2 $\mathbb{P}(A \mid B)$;
    7.3 $\mathbb{P}(\overline{B} \mid A)$.

8. Een boekhouder die het beleid nakijkt van een bedrijf op het vlak van verkopen op krediet stelt vast dat van de 500 klanten er 20% een balans hebben van meer dan 750 €, de anderen hebben 750 € of minder. Qua duurtijd van de rekeningen hebben 30% van de rekeningen meer dan 3 maanden open gestaan, de andere minder. Van degene die meer dan drie maanden open stonden waren er 20% met een balans van meer dan 750 €. We kunnen dit in tabelvorm weergeven:

		A	A'	
		> 750 €	≤ 750 €	Totaal
B	≤ 3 maand	70	280	350
B'	> 3 maand	30	120	150
	Totaal	100	400	500

Oefeningen: vervolg oef 8

1. Hoeveel elementen bevat Ω?
2. Wat is de kans dat een willekeurig gekozen balans minder is dan 750 € en meer dan 3 maanden open stond? Bepaal deze kans, gebruikmakend van voorwaardelijke kansen.
3. Ga na dat A en B' onafhankelijke gebeurtenissen zijn.
4. Bereken $\mathbb{P}(A \cup B)$ op twee manieren, eerst uit de tabelwaarden en dan met de algemene somregel.

De regel van Bayes: Voorbeeld

Een virusscanner ontdekt een bepaald virus. De virusscanner is 95% betrouwbaar voor besmette computers, en voor een niet-besmette computer geeft de virusscanner in 1% van de gevallen aan dat deze toch besmet zou zijn.
Wereldwijd is 4% van de computers besmet met het virus.
Wat is de waarschijnlijkheid dat een computer besmet is met het virus als de virusscanner dit aangeeft?
Laat B de gebeurtenis voorstellen dat de computer besmet is terwijl V de gebeurtenis is dat de virusscanner het virus ontdekt (en dus zegt dat het virus aanwezig is).
We kennen de volgende voorwaardelijke waarschijnlijkheden:

$$\mathbb{P}(V \mid B) = 95\% \quad \text{en} \quad \mathbb{P}(V \mid \overline{B}) = 1\%.$$

Wat gevraagd wordt is

$$\mathbb{P}(B \mid V).$$

De regel van Bayes

Veronderstel dat een gebeurtenis A n verschillende en elkaar wederzijds uitsluitende oorzaken $B_1, B_2, \ldots, B_n$ heeft, dan is

$$\Omega = B_1 \cup B_2 \cup \cdots \cup B_n.$$

Veronderstel bovendien dat je de volgende waarschijnlijkheden kent:

$$\mathbb{P}(A \mid B_i).$$

We vragen ons nu af wat

$$\mathbb{P}(B_j \mid A)$$

is.

De regel van Bayes

Men ziet in dat (WET VAN DE TOTALE WAARSCHIJNLIJKHEID)

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \cap B_i).$$

Er geldt ook

$$\mathbb{P}(B_j \mid A) = \frac{\mathbb{P}(A \cap B_j)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \cap B_j)}{\sum_{i=1}^{n} \mathbb{P}(A \cap B_i)}.$$

Tenslotte geldt

$$\mathbb{P}(A \cap B_j) = \mathbb{P}(B_j)\, \mathbb{P}(A \mid B_j)$$

zodat men de REGEL VAN BAYES vindt:

$$\mathbb{P}(B_j \mid A) = \frac{\mathbb{P}(B_j)\, \mathbb{P}(A \mid B_j)}{\mathbb{P}(A)} = \frac{\mathbb{P}(B_j)\, \mathbb{P}(A \mid B_j)}{\sum_{i=1}^{n} \mathbb{P}(B_i)\, \mathbb{P}(A \mid B_i)}$$

Voorbeeld

Drie machines M1, M2 en M3 worden ingezet voor het vervaardigen van computer chips. De machines verschillen in ouderdom en de ervaring leert dat van de chips die afkomstig zijn van de oudste machine M1 (die instaat voor 20% van de totale productie) 8% niet voldoet. Voor M2 (die 30% voor haar rekening neemt) is dit 5% en voor M3 (die de overige 50% vervaardigt) is dit nog 2%.

Wat is de kans dat, als een defecte chip wordt geconstateerd, deze afkomstig is van M1?

Welk percentage van de totale productie voldoet niet?

Oefeningen

1. In het domein “De Visput” heeft 28% van de wandelaars juist één hond mee. Van deze wandelaars-met-één-hond zijn 45% mannen en 55% vrouwen. Verder blijken mannen de voorkeur te geven aan een grote (macho) hond: 70% van de mannen gaat op stap met een grote hond en slechts 30% heeft een kleinere hond bij. Vrouwen verkiezen een kleinere hond: 75% van de vrouwen gaat op stap met een kleine hond en slechts 25% holt achter een grote hond aan. Als je één kleine hond tegenkomt, bepaal dan de kans dat zijn baas een vrouw is.

Oefeningen

2. Veronderstel dat op een multiple choice examen de volgende situatie zich voordoet. Ofwel weet een student het antwoord met 100% zekerheid, ofwel weet hij het antwoord helemaal niet, en dan gokt hij. Veronderstel dat de kans dat een student het antwoord weet gelijk is aan p, en dat er m mogelijke antwoorden zijn per vraag.
   2.1 Wat is de waarschijnlijkheid dat indien het antwoord op een bepaalde vraag correct is, de student het antwoord ook daadwerkelijk wist?
   2.2 Wat wordt deze waarschijnlijkheid als het aantal antwoorden oneindig wordt?

Oefeningen

3. Een computer programma bestaat uit twee modules. De eerste module bevat een fout met waarschijnlijkheid 0.2; de tweede module is complexer en de waarschijnlijkheid dat deze een fout bevat is 0.4. Fouten in de twee modules zijn onafhankelijk van elkaar. Als er enkel een fout zit in de eerste module dan crasht het programma met waarschijnlijkheid 0.5. Voor de tweede module is dit 0.8. Wanneer beide modules een fout bevatten, dan crasht het programma met 90% waarschijnlijkheid.

   Bij uitvoering crasht het programma. Wat is de waarschijnlijkheid dat beide modules een fout bevatten?

Toevalsveranderlijke: voorbeeld

Bij een kansspel trekt een speler een kaart uit een spel kaarten. Er wordt afgesproken dat hij 1 € krijgt bij het trekken van een boer, 2 € bij het trekken van een dame en 3 € bij het trekken van een heer. In alle andere gevallen krijgt hij niets.
We kunnen dit expliciet maken door een afbeelding X van Ω naar $\mathbb{R}$ te definiëren:

$$X : \Omega \to \mathbb{R} : \omega \mapsto X(\omega) = \begin{cases} 1 & \text{als } \omega \in \{\diamond 11, \clubsuit 11, \heartsuit 11, \spadesuit 11\} \\ 2 & \text{als } \omega \in \{\diamond 12, \clubsuit 12, \heartsuit 12, \spadesuit 12\} \\ 3 & \text{als } \omega \in \{\diamond 13, \clubsuit 13, \heartsuit 13, \spadesuit 13\} \\ 0 & \text{anders.} \end{cases}$$

We hechten m.a.w. een *getalwaarde* aan de verschillende gebeurtenissen.

Toevalsveranderlijke: tweede voorbeeld

Wanneer men een muntstuk tweemaal opgooit dan is

$$\Omega = \{\mathrm{KK}, \mathrm{KM}, \mathrm{MK}, \mathrm{MM}\}.$$

Als $X(\omega)$ het aantal keer kop is voor $\omega \in \Omega$, dan hebben we:

$$X(\mathrm{KK}) = 2, \quad X(\mathrm{KM}) = 1, \quad X(\mathrm{MK}) = 1 \quad \text{en} \quad X(\mathrm{MM}) = 0.$$

We hechten m.a.w. opnieuw een *getalwaarde* aan de verschillende gebeurtenissen.

Toevalsveranderlijke: derde voorbeeld

Als men de wijzer van een rad-van-fortuin vanuit een bepaalde beginpositie (bijvoorbeeld Noord) een draai geeft, dan zal hij deze stilstand een willekeurige hoek aangeven, dus $\Omega = [0, 2\pi[$. We kunnen dan een functie H beschouwen die met elke ω de hoek zelf associeert, dus

$$H : \Omega \to \mathbb{R} : \omega \mapsto H(\omega) = \omega.$$

We kunnen eveneens een tweede functie K beschouwen die met elke ω het *kwadrant* associeert waartoe ω behoort, dus

$$K : \Omega \to \mathbb{R} : \omega \mapsto K(\omega) = \begin{cases} 1 & \text{als } \omega \in [0, \pi/2[\\ 2 & \text{als } \omega \in [\pi/2, \pi[\\ 3 & \text{als } \omega \in [\pi, 3\pi/2[\\ 4 & \text{als } \omega \in [3\pi/2, 2\pi[. \end{cases}$$

Toevalsveranderlijke: definitie

Definitie

Een KANSVARIABELE X *is een afbeelding van* Ω *naar* $\mathbb{R}$. *Deze afbeelding associeert met elke mogelijke uitkomst van een kansexperiment dus een reëel getal.*

Opmerking

- In een meer theoretische benadering zijn er nog (technische) restricties voor deze afbeelding, maar dan gaan we hier nu niet op in.
- Op dit moment is er nog *geen* waarschijnlijkheid geassocieerd met deze kansvariabelen.

Discrete Toevalsveranderlijke: definitie

Definitie

Een kansvariabele X is DISCREET *wanneer X slechts een eindig of aftelbaar oneindig aantal waarden aanneemt. Dit wil zeggen dat*

$$\mathrm{bld}(X) = \{x_1, x_2, \ldots, x_n\} \quad \textit{of} \quad \mathrm{bld}(X) = \{x_1, x_2, \ldots\}.$$

Voorbeeld

- De kansvariabele K uit rad-van-fortuin voorbeeld neemt als waarden 1, 2, 3 en 4 aan. Deze kansvariabele is dus discreet.
- De kansvariabele X voor het kaartspel is eveneens discreet. Het beeld hier is $\{0, 1, 2, 3\}$.

Kansmodel

Stel $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ een kansruimte.
Stel $A(x_i)$ is de gebeurtenis dat X op x_i wordt afgebeeld, i.e.

$$A(x_i) = \{\omega \in \Omega \,|\, X(\omega) = x_i\}.$$

We zeggen dan "de kans dat X gelijk is aan x_i" is

$$\mathbb{P}(X = x_i) = \mathbb{P}\big(A(x_i)\big).$$

We noteren ook $\mathbb{P}(X = x_i)$ ook als $f_X(x_i)$.

Definitie

De functie f_X

$$f_X : \mathbb{R} \to [0,1] : x \mapsto f_X(x) = \mathbb{P}(X = x)$$

is de KANSFUNCTIE *van de discrete toevalsveranderlijke X.*

Kansfunctie: voorbeeld

Voor de kansvariabele horend bij het kaartspel geldt, gebruikmakend van de regel van Laplace:

$$
\begin{aligned}
f_X(3) &= \mathbb{P}(X=3) = \mathbb{P}(\{\diamond 13, \clubsuit 13, \heartsuit 13, \spadesuit 13\}) = \frac{4}{52} = \frac{1}{13}, \\
f_X(2) &= \mathbb{P}(X=2) = \mathbb{P}(\{\diamond 12, \clubsuit 12, \heartsuit 12, \spadesuit 12\}) = \frac{4}{52} = \frac{1}{13}, \\
f_X(1) &= \mathbb{P}(X=1) = \mathbb{P}(\{\diamond 11, \clubsuit 11, \heartsuit 11, \spadesuit 11\}) = \frac{4}{52} = \frac{1}{13}, \\
f_X(0) &= \mathbb{P}(X=0) = \mathbb{P}(\{\diamond 1, \diamond 2, \ldots, \spadesuit 9, \spadesuit 10\}) = \frac{40}{52} = \frac{10}{13}.
\end{aligned}
$$

Opmerking:

- Er geldt: $f_X(0) + f_X(1) + f_X(2) + f_X(3) = 1$.
- Er geldt: $0 \leq f_X(x_i) \leq 1$.

Eigenschappen van kansfunctie

Eigenschap

Een kansfunctie f voor een discrete toevalsveranderlijke voldoet aan de volgende eigenschappen:

1. *Een kansfunctie is steeds positief:* $0 \leq f(x)$.
2. *Een kansfunctie is hoogstens één:* $f(x) \leq 1$.
3. *De som van de waarden $f(x)$ voor alle mogelijke uitkomsten x is één:*

$$\sum_{x \in \mathrm{bld}(X)} f(x) = 1.$$

Kansverdelingsfunctie

Met een (discrete) toevalsveranderlijke X kan men ook een *cumulatieve* kansfunctie associëren.

Definitie

De cumulatieve kansfunctie of KANSVERDELINGSFUNCTIE *van een toevalsveranderlijke X wordt gedefinieerd als*

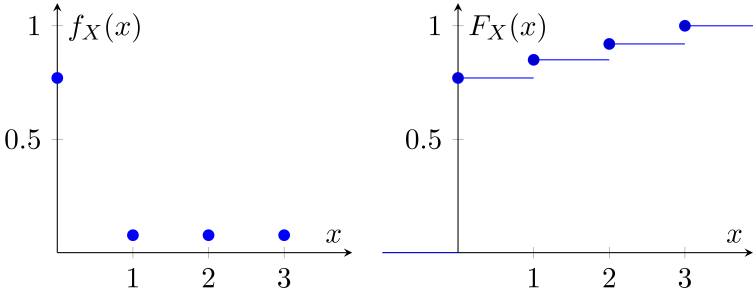
$$\begin{aligned} F_X : \mathbb{R} \to [0,1] : x \mapsto F_X(x) &= \mathbb{P}(X \leq x) \\ &= \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \leq x\}). \end{aligned}$$

Kansverdelingsfunctie: voorbeeld

Voor de kansvariabele X horend bij het kaartspel krijgt men

$$F_X(x) = \begin{cases} 0 & \text{als } x < 0 \\ 10/13 & \text{als } 0 \leq x < 1 \\ 11/13 & \text{als } 1 \leq x < 2 \\ 12/13 & \text{als } 2 \leq x < 3 \\ 1 & \text{als } x \geq 3. \end{cases}$$

Grafieken van kans- en kansverdelingsfunctie



Oefeningen

1. Bij een dobbelspel met één dobbelsteen krijgt men het drievoud uitbetaald van het aantal ogen. Dit geeft aanleiding tot een kansvariabele X. Geef de uitkomstenverzameling Ω, f_X en F_X. Geef de grafiek van zowel f_X als F_X. Men spreekt van de DISCRETE UNIFORME VERDELING .
2. Men gooit een evenwichtig muntstuk op tot voor de eerste keer "munt" verschijnt. We noemen X het daartoe benodigde aantal worpen. Wat is het beeld van X? Geef $f_X(x)$ en $F_X(x)$. Schets hun grafiek. Hoe groot is de kans dat het aantal worpen minstens 3 en ten hoogste 6 bedraagt?
3. Bij een dobbelspel met twee dobbelstenen noteert men X als de som van het aantal ogen. Onderzoek f_X en F_X.
4. Veronderstel dat het munstuk in Voorbeeld 2.2 zodanig vervalst is dat $\mathbb{P}(K) = 6/10$. Bepaal f_X en F_X.

Continue toevalsveranderlijken: inleiding

Bekijk de kansvariabele H uit rad-van-fortuin voorbeeld.

Intuïtief is duidelijk dat de kans dat de wijzer *exact* stopt op een hoek $1/2$ (radiaal) gelijk is aan nul.

Want er zijn (overaftelbaar) oneindig veel mogelijkheden waar de wijzer kan stoppen, en elk heeft evenveel “kans”.

Wel zinvol is de kans dat de wijzer stopt *tussen* twee hoeken a en b.

Kansverdelingsfunctie

Definitie

De KANSVERDELINGSFUNCTIE F_X *van een toevalsveranderlijke* X *is de functie van* $\mathbb{R}$ *naar het het interval* $[0,1]$ *gegeven door:*

$$F_X : \mathbb{R} \to [0,1] : x \mapsto F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \leq x\})$$

Definitie

Een toevalsveranderlijke X *is* CONTINU *als er een functie* f_X *van* $\mathbb{R}$ *naar* $\mathbb{R}^+$ *bestaat zodanig dat*

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\,\mathrm{d}y.$$

De functie f_X *wordt de* KANSDICHTHEID *genoemd.*

Grafisch

Zie bord!

Eigenschap

De functie f_X voldoet aan de volgende eigenschappen.

1. *De functie f_X is nergens negatief, i.e. $f_X(x) \geq 0$.*
2. *Als men f_X integreert over $\mathbb{R}$ dan bekomt men als uitkomst 1:*
$$\int_{-\infty}^{+\infty} f_X(x)\,\mathrm{d}x = 1.$$
3. *De kans dat X behoort tot een interval $[a, b]$ wordt uitgedrukt d.m.v. de volgende integraal:*
$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)\,\mathrm{d}x.$$

Opmerkingen:

- Er bestaan *continue* toevalsveranderlijk waarvoor $f_X(x) > 1$ voor bepaalde reële getallen x.
- $f_X(a)$ is *niet* de kans dat de toevalsveranderlijke X de waarde a aanneemt. Deze waarschijnlijkheid is immers steeds nul:

$$\mathbb{P}(X = a) = \mathbb{P}(a \leq X \leq a) = \int_a^a f_X(x)\,\mathrm{d}x = 0.$$

- Voor een *continue* toevalsveranderlijke X geldt dat

$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a < X \leq b) = \mathbb{P}(a < X < b).$$

Dit is *niet* geldig voor discrete toevalsveranderlijken.

Voorbeeld

We beschouwen de kansvariabele H uit het rad-van-fortuin experiment.
Uit symmetrie-overwegingen geldt:

$$\mathbb{P}([a,b]) = \frac{b-a}{2\pi}.$$

(Hoe werd $1/2\pi$ bekomen?)
Hieruit volgt dat (voor $0 \leq x < 2\pi$)

$$\begin{aligned} F_H(x) &= \mathbb{P}(H \leq x) \\ &= \mathbb{P}(0 \leq H \leq x) \\ &= \mathbb{P}(\{\omega \in \Omega \,|\, 0 \leq H(\omega) \leq x\}) \\ &= \mathbb{P}(\{\omega \in \Omega \,|\, 0 \leq \omega \leq x\}) \\ &= \mathbb{P}([0,x]) \\ &= \frac{x}{2\pi}. \end{aligned}$$

Voorbeeld: vervolg

Samengevat:

$$F_H(x) = \begin{cases} 0 & \text{als } x < 0 \\ \dfrac{x}{2\pi} & \text{als } 0 \leq x < 2\pi \\ 1 & \text{als } x \geq 2\pi. \end{cases}$$

Hieruit volgt dan:

$$f_H(x) = \begin{cases} \dfrac{1}{2\pi} & \text{als } 0 \leq x < 2\pi \\ 0 & \text{anders.} \end{cases}$$

We kunnen deze functies ook grafisch voorstellen.

Discrete kansvariabele: verwachtingswaarde

- Stel X een discrete kansvariabele, i.e. $X\colon \Omega \to \mathbb{R}\colon \omega \mapsto X(\omega)$.
- Voer een groot aantal, N, onafhankelijke herhalingen van het experiment uit.
- Men bekomt een reeks van reële getallen $X(\omega_i)$ waarbij ω_i de uitkomst van het i-de experiment voorstelt.
- Bereken het (rekenkundig) gemiddelde van deze getallen:

$$\frac{1}{N}\big(X(\omega_1) + X(\omega_2) + \cdots + X(\omega_N)\big).$$

- Groepeer volgens uitkomst x_i $(i = 1\ldots, n)$

$$\begin{aligned}&\frac{1}{N}\big(x_1 \cdot \#x_1 + x_2 \cdot \#x_2 + \cdots + x_n \cdot \#x_n\big)\\ &\qquad = x_1 \cdot \frac{\#x_1}{N} + x_2 \cdot \frac{\#x_2}{N} + \cdots + x_n \cdot \frac{\#x_n}{N}.\end{aligned}$$

- Nu zal $\#x_i/N$ naderen naar $\mathbb{P}(X = x_i)$.

Discrete kansvariabele: verwachtingswaarde

Definitie

De VERWACHTINGSWAARDE *van een discrete toevalsveranderlijke X wordt genoteerd als μ_X of $\mathrm{E}(X)$, en is een gewogen gemiddelde van de waarden x_i die X kan aannemen met de respectievelijke kansen als gewichten. In formulevorm:*

$$\mathrm{E}(X) = \sum_i x_i \, \mathbb{P}(X = x_i) = \sum_i x_i f_X(x_i).$$

Voorbeeld

Beschouw de toevalsveranderlijke voor het kaartspel.

$$f_X(3) = \mathbb{P}(X = 3) = \frac{1}{13}, \qquad f_X(2) = \mathbb{P}(X = 2) = \frac{1}{13},$$
$$f_X(1) = \mathbb{P}(X = 1) = \frac{1}{13}, \qquad f_X(0) = \mathbb{P}(X = 0) = \frac{10}{13}.$$

We berekenen de verwachte opbrengst

$$\begin{aligned}
\mu_X &= \sum_i x_i f_X(x_i) \\
&= 0 \cdot f_X(0) + 1 \cdot f_X(1) + 2 \cdot f_X(2) + 3 \cdot f_X(3) \\
&= 0 \cdot \frac{10}{13} + 1 \cdot \frac{1}{13} + 2 \cdot \frac{1}{13} + 3 \cdot \frac{1}{13} = \frac{6}{13}.
\end{aligned}$$

Een spel is EERLIJK als de verwachte opbrengst gelijk is aan de inzet die men moet doen.

Variantie

Definitie

De VARIANTIE *van een toevalsveranderlijke X, genoteerd σ_X^2, is de gewogen gemiddelde kwadratische afwijking t.o.v. zijn verwachtingswaarde. In symbolen:*

$$\sigma_X^2 = \sum_i (x_i - \mu_X)^2 \, \mathbb{P}(X = x_i) = \sum_i (x_i - \mu_X)^2 f_X(x_i).$$

Opmerking:

- De variantie drukt uit hoe "gespreid" de grafiek van de kansfunctie van de kansvariabele X is.
- Hoe kleiner de variantie hoe meer geconcentreerd de kansfunctie is rond het gemiddelde.
- De STANDAARDAFWIJKING, genoteerd σ_X, is de vierkantswortel uit de variantie.

Voorbeeld

We berekenen de variantie σ_X^2 voor de toevalsveranderlijke X van het kaartspel. We weten reeds dat $\mu_X = 6/13$, en we vinden dus

$$\begin{aligned}
\sigma_X^2 &= \sum_i (x_i - \mu_X)^2 f_X(x_i) \\
&= (0 - \frac{6}{13})^2 \cdot f_X(0) + (1 - \frac{6}{13})^2 \cdot f_X(1) + (2 - \frac{6}{13})^2 \cdot f_X(2) \\
&\qquad + (3 - \frac{6}{13})^2 \cdot f_X(3) \\
&= \frac{36}{169} \cdot \frac{10}{13} + \frac{49}{169} \cdot \frac{1}{13} + \frac{400}{169} \cdot \frac{1}{13} + \frac{1089}{169} \cdot \frac{1}{13} \\
&= \frac{146}{169}.
\end{aligned}$$

Tweede voorbeeld: intuïtie variantie

Beschouw twee toevalsveranderlijken X en Y. De toevalsveranderlijke X stelt het aantal ontvangen emails per dag voor van een eerste gebruiker. Deze gebruiker ontvangt elke dag ofwel 48 ofwel 52 emails, elk met 50% kans. De toevalsveranderlijke Y stelt het aantal ontvangen emails voor van een tweede gebruiker. Deze tweede gebruiker ontvangt met 50% kans geen enkele email, en met 50% kans 100 emails.

- Stel de kansfuncties op van X en Y.
- Bereken de verwachtingswaarde van X en Y. Wat merk je?
- Wie ervaart er een grotere variabiliteit m.b.t. het aantal emails in zijn inbox?
- Bereken de variantie van X en Y.

Continue toevalsveranderlijke

Definitie

De verwachtingswaarde $\mathrm{E}(X)$ *(of* μ_X*) van een continue toevalsveranderlijke* X *wordt gedefinieerd als:*

$$\mu_X = \int_{-\infty}^{+\infty} x\, f_X(x)\, \mathrm{d}x,$$

terwijl de variantie σ_X^2 *wordt gegeven door*

$$\sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f_X(x)\, \mathrm{d}x.$$

Opmerking:

Men bekomt deze formules door in de formules voor discrete kansvariabelen de som te vervangen door een integratie.

Voorbeeld

We berekenen μ_H voor de kansvariabele H uit het rad-van-fortuin.

$$\begin{aligned}
\mu_H &= \int_{-\infty}^{+\infty} x f_H(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{0} x f_H(x)\,\mathrm{d}x + \int_{0}^{2\pi} x f_H(x)\,\mathrm{d}x + \int_{2\pi}^{+\infty} x f_H(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{0} x \cdot 0\,\mathrm{d}x + \int_{0}^{2\pi} x \frac{1}{2\pi}\,\mathrm{d}x + \int_{2\pi}^{+\infty} x \cdot 0\,\mathrm{d}x \\
&= \frac{1}{2\pi} \int_{0}^{2\pi} x\,\mathrm{d}x \\
&= \pi.
\end{aligned}$$

Opmerking:

- Waarde integraal is gelijk aan oppervlakte driehoek.
- Is dit wat je zou verwachten als gemiddelde?
- Variantie iets moeilijker te berekenen als oppervlakte.

Law of the unconscious statistician

- Stel X een toevalsveranderlijke, i.e. $X\colon \Omega \to \mathbb{R}$;
- stel g een reële functie is, i.e. $g\colon \mathbb{R} \to \mathbb{R}$
- dan is $Y = g \circ X$ ook een toevalsveranderlijke, de formule voor $g \circ X$ is:

$$g \circ X\colon \Omega \to \mathbb{R}\colon \omega \mapsto g(X(\omega)).$$

We kunnen voor deze nieuwe toevalsveranderlijke $Y = g(X)$ de verwachtingswaarde $\mathrm{E}(Y)$ berekenen *zonder* dat we expliciet de kansfunctie f_Y moeten opstellen!

Law of the unconscious statistician

Eigenschap

Als X een discrete toevalsveranderlijke is en g is een functie van $\mathbb{R}$ naar $\mathbb{R}$ dan geldt:

$$\mathrm{E}(Y) = \mathrm{E}(g(X)) = \sum_{x \in bld(X)} g(x) f_X(x) = \sum_{x \in bld(X)} g(x)\, \mathbb{P}(X = x).$$

Merk op dat in deze formule de kansfunctie van X wordt gebruikt en niet die van Y.

Deze eigenschap is ook geldig voor continue toevalsveranderlijken maar met integralen i.p.v. sommaties.

Voorbeeld

Veronderstel dat X de waarden -2, -1, 1 en 3 aanneemt met waarschijnlijkheid $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{4}$ en $\frac{3}{8}$ respectievelijk.
Veronderstel dat g de functie is die x afbeeldt op x^2.

Bereken $E(Y)$ op twee manieren en ga na dat de uitkomst in beide gevallen hetzelfde is.

Eigenschappen

Eigenschap

De verwachtingswaarde en variantie van een (discrete of continue) toevalsveranderlijke X voldoen aan de volgende eigenschappen:

1. *Als X constant is, i.e. $X(\omega) = k$ voor alle elementen ω van Ω, dan is $\mathrm{E}(X) = k$.*
2. *Als $a \in \mathbb{R}$ een constante is, dan geldt:*

$$\mathrm{E}(X + a) = \mathrm{E}(X) + a,$$

   waaruit in het bijzonder volgt dat

$$\mathrm{E}(X - \mu_X) = 0.$$

3. *Als $a \in \mathbb{R}$ een constante is, dan geldt:*

$$\mathrm{E}(aX) = a\,\mathrm{E}(X).$$

Eigenschappen: vervolg

Eigenschap

4. *Er geldt steeds dat*

$$\sigma_X^2 = \mathrm{E}(X^2) - \mu_X^2.$$

   Deze formule geeft de mogelijkheid om de variantie efficiënter te berekenen dan rechtstreeks via de definitie.

5. *De variantie wijzigt niet als we de toevalsveranderlijke "verschuiven". Er geldt voor elke constante $a \in \mathbb{R}$ dat*

$$\sigma_{(X+a)}^2 = \sigma_X^2.$$

6. *Vermenigvuldigen met een constante wijzigt de variantie op een kwadratische manier:*

$$\sigma_{aX}^2 = a^2\sigma_X^2.$$

Voorbeeld

- We berekenen de variantie van de kansvariabele voor het kaartspel op een andere manier.
- We starten met de berekening van $\mathrm{E}(X^2)$:

$$\mathrm{E}(X^2) = \sum_i x_i^2 f_X(x_i) = 0^2 \frac{10}{13} + 1^2 \frac{1}{13} + 2^2 \frac{1}{13} + 3^2 \frac{1}{13} = \frac{14}{13}.$$

- Dus wordt de variantie volgens (4) gegeven door

$$\sigma_X^2 = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{14}{13} - \frac{6^2}{13^2} = \frac{146}{169}.$$

Dit is in overeenstemming met hetgeen vroeger werd bekomen.

Voorbeeld

Veronderstel dat we de opbrengst van het kaartspel *verdubbelen*, i.e. we creëren een toevalsveranderlijke $Y = 2X$. Expliciet wordt Y dus gegeven door:

$$Y : \Omega \to \mathbb{R} : \omega \mapsto Y(\omega) = \begin{cases} 2 & \text{als } \omega \in \{\diamond 11, \clubsuit 11, \heartsuit 11, \spadesuit 11\} \\ 4 & \text{als } \omega \in \{\diamond 12, \clubsuit 12, \heartsuit 12, \spadesuit 12\} \\ 6 & \text{als } \omega \in \{\diamond 13, \clubsuit 13, \heartsuit 13, \spadesuit 13\} \\ 0 & \text{anders.} \end{cases}$$

Als we nu $\mathrm{E}(Y)$ berekenen, dan vinden we:

$$\begin{aligned} \mu_Y &= \sum_i x_i f_Y(x_i) = 0 \cdot f_Y(0) + 2 \cdot f_Y(2) + 4 \cdot f_Y(4) + 6 \cdot f_Y(6) \\ &= 0 \cdot \frac{10}{13} + 2 \cdot \frac{1}{13} + 4 \cdot \frac{1}{13} + 6 \cdot \frac{1}{13} = \frac{12}{13}. \end{aligned}$$

Merk op dat dus inderdaad $\mathrm{E}(Y) = \mathrm{E}(2X) = 2\,\mathrm{E}(X)$.

Voorbeeld: vervolg

- We berekenen nu nog de variantie σ_Y^2 van Y:

$$
\begin{aligned}
\sigma_Y^2 &= \sum_i (x_i - \mu_Y)^2 f_Y(x_i) \\
&= (0 - \frac{12}{13})^2 \cdot f_Y(0) + (2 - \frac{12}{13})^2 \cdot f_Y(2) \\
&\quad + (4 - \frac{12}{13})^2 \cdot f_Y(4) + (6 - \frac{12}{13})^2 \cdot f_Y(6) \\
&= \frac{584}{169}.
\end{aligned}
$$

- We vinden dus dat $\sigma_Y^2 = \sigma_{2X}^2 = 4\sigma_X^2$, in overeenstemming met puntje (6) van de eigenschap.

Gevolg: standaardisatie toevalsveranderlijke

Eigenschap

Voor elke (niet-constante) toevalsveranderlijke X geldt dat de toevalsveranderlijke Z, met

$$Z = \frac{X - \mu_X}{\sigma_X}$$

gemiddelde nul en variantie één heeft, i.e.

$$\mu_Z = 0 \quad \textit{en} \quad \sigma_Z^2 = 1.$$

Oefeningen

1. Bij een kleine verzekeringsmaatschappij heeft men ondervonden dat – gemiddeld – 1 op 10000 polissen resulteert in een schadeclaim van 200000 €, 1 op 1000 in een claim van 50000 €, 1 op 50 in een claim van 2000 € en de rest in een schadeclaim van 0 €. Men wenst nu te weten (teneinde een prijsaanpassing door te voeren) wat men gemiddeld uitbetaalt per polis.

Oefeningen

2. Een kioskhouder verkoopt elke zaterdag een aantal exemplaren van een bepaalde krant. Het aantal exemplaren dat op een willekeurige zaterdag wordt *gevraagd* kan beschouwd worden als een kansvariabele. Nadat hij de verkoop een tijdlang heeft opgevolgd bekomt hij volgende resultaten

aantal	kans	aantal	kans	aantal	kans
30	0.10	40	0.30	50	0.05
35	0.35	45	0.20		

2.1 De kioskhouder koopt elke zaterdag 40 kranten in. Hoe groot is de kans dat hij niet aan de vraag kan voldoen?
2.2 Wat is de verwachtingswaarde van de vraag op een willekeurige zaterdag? Wat is de variantie van de vraag?
2.3 De kranten kosten de kioskhouder 0.60 €, terwijl de verkoopprijs 1.15 € bedraagt. Kranten die niet verkocht werden moeten worden weggegooid. Bereken de verwachte winst van de kioskhouder op een willekeurige zaterdag, als hij 40 kranten inkoopt.

Oefeningen

3. Een fabriek heeft twee afdelingen. Tot nu toe werden reparaties uitgevoerd in twee afzonderlijke werkplaatsen. Het bedrijf overweegt nu een nieuwe gemeenschappelijke werkplaats in te richten. Het machinepark werd gedurende 100 dagen bestudeerd en men vond volgende resultaten:

Defect per dag	0	1	2
Afdeling A	30	50	20
Afdeling B	20	30	50

Men neemt aan dat de defecten onafhankelijk zijn. Hoeveel defecten mag men per dag verwachten in de gemeenschappelijke werkplaats? Neem aan dat men uit de tabel de exacte kansfuncties kan afleiden.

Oefeningen

4. Veronderstel dat men in een ziekenhuis bloedtesten uitvoert voor een bepaalde ziekte waaraan ongeveer 1 op 100 mensen lijden. De mensen komen naar het ziekenhuis in groepen van 50 (bijvoorbeeld scholen) en de directeur vraagt zich af of – in plaats van individueel te testen – het niet beter zou zijn de 50 personen samen te testen. Als ze allemaal negatief zijn kan hij de hele groep gezond verklaren, zoniet zou hij overgaan tot individuele testen.
   4.1 Wat is het verwachte aantal testen als individueel wordt gescreend?
   4.2 Wat is het verwachte aantal testen als in groep wordt gescreend?
   4.3 Wat zou jij nu beslissen als directeur?

Oefeningen

5. Bij het roulettespel kan men een inzet doen op 37 nummers, genummerd van 0 tot 36. Er zijn 18 rode vakjes, 18 zwarte vakjes en het nummer 0 is wit. Als men een bedrag inzet op een bepaald nummer en dit nummer wint, dan krijgt men 36 keer het oorspronkelijke bedrag uitbetaald. Is dit een eerlijk spel?

Inleiding

Vaak is het handiger te werken met *theoretische* verdelingen dan met empirische gegevens.
Beschouw bijvoorbeeld volgende empirische verdeling voor de uitkomsten bij het 60000 keer gooien van een dobbelsteen:

Ogen	h_i
1	9978
2	10021
3	10013
4	9903
5	9986
6	10099

Het is duidelijk dat de (theoretische) uniforme verdeling hier beter is dan de experimentele.

We beschouwen nu enkele veelgebruikte theoretische kansverdelingen die handig kunnen gebruikt worden om veel praktische situaties te modelleren.

Bernoulliverdeling

Veel eenvoudige experimenten hebben slechts twee mogelijke uitkomsten. Men noemt ze BERNOULLI-EXPERIMENTEN.

De volgende experimenten zijn voorbeelden van Bernoulli-experimenten:

- het opgooien van een muntstuk: kop of munt;
- een dobbelsteen gooien: zes of geen zes;
- de geboorte van een kind bij een bepaald ouderpaar: meisje of jongen;
- een bal uit een vaas met witte en zwarte ballen trekken: wit of zwart.

Kansfunctie

- De ene uitkomst "succes" wordt voorgesteld door 1; andere "mislukking" wordt voorgesteld door 0.
- Kans op succes: p, kans op mislukking $q = 1 - p$.
- Kansfunctie is dus:

x_i	0	1
$\mathbb{P}(X = x_i)$	q	p

Eigenschap

Als X een toevalsveranderlijke is die een Bernoulliverdeling volgt met kans op succes p, dan geldt

$$\mu_X = p \quad \textit{en} \quad \sigma_X^2 = p(1-p).$$

Binomiale verdeling: aantal successen

- De toevalsveranderlijke X die **telt** hoeveel successen er voorkomen wanneer men n onafhankelijke Bernoulli-experimenten uitvoert volgt een BINOMIALE VERDELING.
- We noteren
$$X \sim B(n, p).$$

Eigenschap

Veronderstel dat $X \sim B(n, p)$, dan wordt de kansfunctie van X gegeven door:

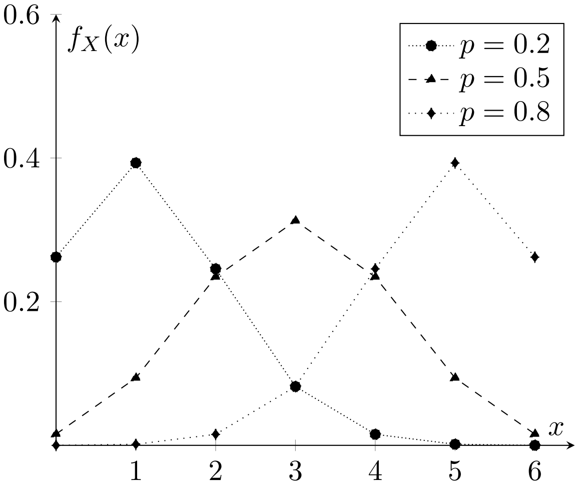
$$f_X(k) = \mathbb{P}(X = k) = C_n^k p^k q^{n-k} = \binom{n}{k} p^k q^{n-k}, \quad \textit{als } 0 \leq k \leq n.$$

Voorbeelden

De volgende toevalsveranderlijken volgen een binomiale verdeling:

- De toevalsveranderlijke die het aantal keer kop telt wanneer men 5 keer met een muntstuk gooit.
- De toevalsveranderlijke die het aantal winnende loten telt onder 10 gekochte lotjes.
- De toevalsveranderlijke die het aantal jongens telt onder 4 geboorten bij een bepaald ouderpaar.
- De toevalsveranderlijke die het aantal witte ballen telt wanneer men drie ballen uit een vaas met witte en zwarte ballen trekt. De getrokken bal wordt teruggelegd.

Enkele grafieken



Verwachtingswaarde en variantie

Eigenschap

Als $X \sim B(n, p)$ *dan geldt*

$$\mu_X = np$$

en

$$\sigma_X^2 = np(1 - p).$$

Opmerking:

- Logische uitbreiding van de formules voor Bernouilliverdeling.

Kansverdelingsfunctie

- De kansverdelingsfunctie F_X geeft de kans op *ten hoogste* k successen: $F_X(k) = \mathbb{P}(X \leq k)$.
- Met de complementregel vindt men dan gemakkelijk

$$\mathbb{P}(X \geq k) = 1 - \mathbb{P}(X < k) = 1 - \mathbb{P}(X \leq k-1) = 1 - F(k-1).$$

- Hierbij hebben we gebruikgemaakt van het feit dat de binomiale verdeling een *discrete* verdeling is.
- Dit kan soms helpen om het rekenwerk te verkorten.

Geometrische verdeling: wachten op het eerste succes

Veronderstel dat we onafhankelijke Bernoulli-experimenten uitvoeren met kans op succes p. Na hoeveel uitvoeringen bekomt men voor het eerst een succes?

De kansveranderlijke X die het **rangnummer** aangeeft van het eerste succes volgt een GEOMETRISCHE verdeling.

Voorbeelden:

- De kansveranderlijke die telt hoe vaak men met een muntstuk moet werpen alvorens voor de eerste maal kop te bekomen.
- De kansveranderlijke die aangeeft hoeveel bloedafnames men moet doen alvorens men de bloedgroep AB aantreft.
- Het aantal keer dat men op de roulette moet spelen alvorens men voor de eerste maal wint.

Kansfunctie

Eigenschap

Als X geometrisch verdeeld is met parameter p, dan wordt de kansfunctie f_X gegeven door

$$f_X(k) = \mathbb{P}(X = k) = q^{k-1}p, \quad \textit{voor } k \geq 1.$$

Opmerking:

- De reekssom van de meetkundige (ofte geometrische) reeks met reden q, is:

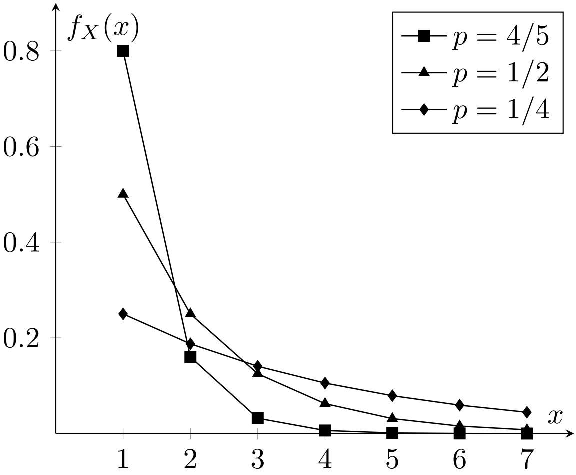
$$1 + q + q^2 + q^3 + \cdots = \sum_{i=0}^{\infty} q^i = \frac{1}{1-q}.$$

- Dus is

$$\sum_{k=1}^{\infty} q^{k-1}p = p\sum_{k=1}^{\infty} q^{k-1} = p\frac{1}{1-q} = 1.$$

- Dus: ooit zal er een succes voorkomen (als $p > 0$).

Kansfunctie: grafisch



Verwachtingswaarde en variantie

Eigenschap

Als X geometrisch verdeeld is met parameter p dan geldt:

$$\mu_X = \frac{1}{p}$$

en

$$\sigma_X^2 = \frac{q}{p^2}.$$

Opmerking:

De gemiddelde "wachttijd"[1] is dus omgekeerd evenredig met de kans op succes. Logisch!

[1] De geometrische verdeling is uiteraard geen continue verdeling!

Markov-eigenschap

De geometrische verdeling bezit de *Markov-eigenschap*; ze heeft geen geheugen.

Meer formeel:

$$\mathbb{P}(X > m + n \,|\, X > m) = \mathbb{P}(X > n).$$

Er geldt (waarom?):

$$\mathbb{P}(X > k) = q^k.$$

Nu dient men enkel nog de definitie van voorwaardelijke kans toe te passen.

Toepassing

- Via de geometrische verdeling kan bepaald worden hoeveel trekkingen (met teruglegging) men moet uitvoeren om uit een populatie van n verschillende elementen er juist r verschillende te verzamelen.
- Stel dat X_k het aantal trekkingen tussen het bekomen van de eerste k elementen en het $(k+1)$-ste verschillende element aanduidt.
- X_k volgt een geometrische verdeling met parameter $(n-k)/n$.
- Dus, het aantal benodigde trekkingen om de eerste r verschillende elementen te bekomen is dan te schrijven als

$$1 + X_1 + X_2 + \cdots + X_{r-1}.$$

Voorbeeld

We bepalen de gemiddelde wachttijd wanneer we uit 6 verschillende elementen er 3 verschillende wensen te bekomen. Het aantal benodigde trekkingen is verdeeld volgens

$$1 + X_1 + X_2,$$

met X_i geometrisch verdeeld is met $p = (6 - i)/6$ voor $i \in \{1, 2\}$. Er geldt

$$\mathrm{E}(X_i) = \frac{6}{6 - i},$$

en dus is

$$\mathrm{E}(1 + X_1 + X_2) = 1 + \mathrm{E}(X_1) + \mathrm{E}(X_2) = 1 + \frac{6}{5} + \frac{6}{4} = \frac{37}{10}.$$

Men moet dus gemiddeld 3.7 trekkingen doen alvorens 3 verschillende elementen werden bekomen bij trekking met teruglegging uit een populatie van 6 elementen.

Poisson verdeling: zeldzame gebeurtenissen

- Stel een *zeer groot* aantal, genoteerd n, onafhankelijke uitvoeringen van eenzelfde Bernoulli-experiment.
- Stel de kans op succes p *zeer klein* is.
- Stel dat het gemiddeld aantal successen $n \cdot p$ *ongeveer constant* is.
- Deze constante wordt genoteerd met λ.
- Het aantal successen voldoet dan aan wat men noemt een $\textsc{Poisson}$-verdeling met parameter λ.

Voorbeelden

1. Het aantal auto's dat per dag stopt aan een bepaald benzinestation om te tanken. In dit geval is n het aantal auto's dat per dag aan dat benzinestation passeert, terwijl p de kans is dat een auto stopt om te tanken.
2. Het aantal foutieve transmissietekens per seconde bij satellietcommunicatie. Hier is n het aantal tekens dat per seconde overgeseind wordt via de satelliet en p is de kans dat een fout optreedt.
3. Het aantal complicaties per dag bij injectie van een bepaald medicijn. Hier is n het aantal injecties met een bepaald medicijn per dag terwijl p de kans is dat een bepaalde schadelijke nevenreactie optreedt.

Kansfunctie

Eigenschap

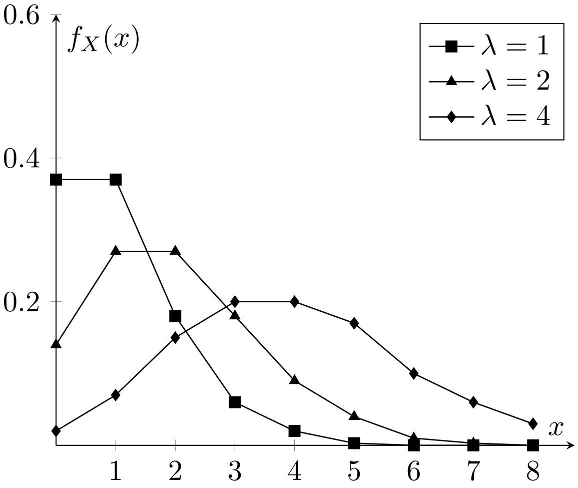
Als X Poisson-verdeeld is met parameter λ, dan wordt zijn kansfunctie gegeven door

$$f_X(k) = \mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad \textit{met} \quad k \geq 0.$$

Opmerking:

Het bewijs volgt door een limietovergang van de kansfunctie van een binomiaal verdeelde toevalsveranderlijke.

Kansfunctie: grafisch



Vergelijking met binomiale verdeling

100 koppels (i, j) met $1 \leq i, j \leq 10$. Tel aantal voorkomens $(7, 7)$.

k	binomiaal	Poisson	empirisch
0	0,366032	0,367879	0,41
1	0,369730	0,367879	0,34
2	0,184865	0,183940	0,16
3	0,060999	0,061313	0,08
4	0,014942	0,015328	0,00
5	0,002898	0,003066	0,01
6	0,000463	0,000511	0,00
7	0,000063	0,000073	0,00
8	0,000007	0,000009	0,00
9	0,000001	0,000001	0,00

Verwachtingswaarde en variantie

Eigenschap

Als de toevalsveranderlijke X Poisson-verdeeld is met parameter λ, dan geldt

$$\mu_X = \lambda$$

en

$$\sigma_X^2 = \lambda.$$

Opmerking:

Dit is in overeenstemming met hetgeen we reeds weten voor de binomiale verdeling! (Laat $n \to \infty$ en $p \to 0$ met $n \cdot p = \lambda$.)

Willekeurig tijdsinterval

- Vaak wordt het eenheidstijdsinterval vervangen door een willekeurig tijdsinterval $[0, t]$. Dit leidt tot de kans

$$\mathbb{P}(X = k) = e^{-\lambda t} \cdot \frac{(\lambda t)^k}{k!}$$

  voor het vinden van precies k successen in het interval $[0, t]$.
- De kans op geen enkel succes in het interval $[0, t]$ is dan

$$\mathbb{P}(X = 0) = e^{-\lambda t}$$

- De kans op 1 of meer successen

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - e^{-\lambda t}.$$

- De parameter λ is een fysische constante die de *dichtheid* van het aantal successen aangeeft.

Oefeningen

1. Een verzekeringsmaatschappij sluit met 10 personen een levensverzekering af. De kans dat een individuele persoon nog in leven is na 30 jaar is 60%. Bereken de kans dat na 30 jaar
   1.1 alle verzekerden nog leven;
   1.2 minstens 3 personen nog in leven zijn;
   1.3 exact 4 personen in leven zijn.
2. Om de verkoop van een bepaald artikel te stimuleren heeft een fabrikant in een kwart van de pakjes een geschenkje verpakt. Iemand koopt twee pakjes.
   2.1 Bereken de kans op twee geschenken.
   2.2 Wat is het gemiddeld aantal geschenken dat zo'n persoon mag verwachten, i.e. wat is μ?
   2.3 Bepaal de variantie σ^2.

Oefeningen

3. Een wijnhandelaar blijkt zijn flessen enigszins onnauwkeurig te vullen. Hierdoor voldoet 10% van de afgeleverde flessen niet aan de inhoudsnorm van het etiket.
   3.1 Een consument koopt 12 flessen wijn. Hoe groot is de kans dat precies 2 flessen niet aan de norm voldoen?
   3.2 Iemand koopt 144 flessen. Hoe groot is de kans dat hoogstens 10 flessen niet aan de norm voldoen?
4. Bij een injectie met een bepaald medicijn is de kans op een schadelijke nevenreactie 0,001. Bereken de kans dat bij 2000 inspuitingen
   4.1 3 patiënten deze reactie vertonen;
   4.2 meer dan 5 patiënten de reactie vertonen.

Oefeningen

5. Het aantal klanten dat per minuut een postkantoor binnenkomt, mag worden beschouwd als een kansvariabele met een Poisson-verdeling. De kans dat binnen een minuut niemand binnenkomt, bedraagt 0,6065. Bepaal de kans dat in een periode van 10 minuten er meer dan 10 klanten binnenkomen.
6. Bij een telefooncentrale komen gemiddeld 180 oproepen per uur binnen. Het aantal oproepen per uur mag beschouwd worden als een toevalsveranderlijke met een Poisson-verdeling. In een minuut kunnen hoogstens 6 gesprekken verwerkt worden. Bereken de kans dat in een bepaalde minuut overbelasting optreedt.
7. Op een bepaalde verkeersweg tussen twee steden gebeuren per jaar gemiddeld 10 verkeersongevallen. Bereken de kans dat er tijdens het komende jaar 8 of meer ongevallen gebeuren in de veronderstelling dat het aantal verkeersongevallen per jaar een Poisson-verdeling volgt.

Oefeningen

8. Een webrobot zoekt naar een bepaald sleutelwoord in een reeks van onafhankelijk gekozen webpagina's. Men gaat ervan uit dat 15% van de webpagina's dit sleutelwoord bevatten.
   8.1 Wat is de kans dat de webrobot een webpagina vindt die het sleutelwoord bevat in de eerste vijf bezochte websites?
   8.2 Hoeveel webpagina's moet de webrobot gemiddeld bezoeken om een webpagina te vinden die het sleutelwoord bevat?

Uniforme verdeling

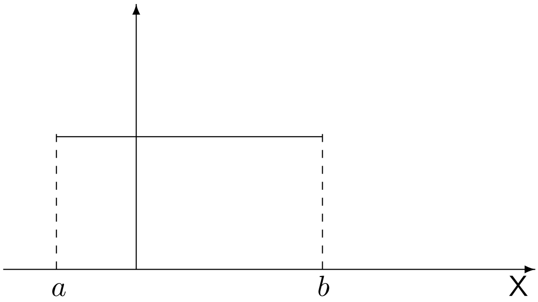
Boven een bepaalde *ondergrens* a en onder een bepaalde *bovengrens* b is de kansdichtheid constant (en verschillend van nul).

Eigenschap

Als X een continue UNIFORME VERDELING *volgt met grenzen a en b, dan wordt de kansdichtheid van X gegeven door*

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & \textit{als } a \leq x \leq b, \\ 0 & \textit{anders}. \end{cases}$$

Kansdichtheid: grafisch



Verwachtingswaarde en variantie

Eigenschap

Als X een continue uniforme verdeling volgt met grenzen a en b dan geldt

$$\mu_X = \frac{(a+b)}{2},$$

en

$$\sigma_X^2 = \frac{(b-a)^2}{12}.$$

Opmerkingen:

- Men kan dit bewijzen door (eenvoudige) integralen te berekenen.
- De formule voor μ_X zou, gezien de *symmetrie* van de verdeling, intuïtief duidelijk moeten zijn.
- Hoe groter $b-a$, hoe groter de variantie.

De exponentiële verdeling

- Start met Poisson-proces met parameter λ.
- Wat is de **tijd** alvorens het eerste succes voorkomt?
- Deze tijd is een toevalsveranderlijke, en volgt een EXPONENTIËLE VERDELING.

Kansdichtheid en kansverdelingsfunctie

Eigenschap

De kansdichtheid en kansverdelingsfunctie van een exponentieel verdeelde veranderlijke T met parameter λ worden respectievelijk gegeven door:

$$f_T(t) = \begin{cases} 0 & \textit{als } t < 0 \\ \lambda e^{-\lambda t} & \textit{als } t \geq 0, \end{cases}$$

en

$$F_T(t) = \begin{cases} 0 & \textit{als } t < 0 \\ 1 - e^{-\lambda t} & \textit{als } t \geq 0. \end{cases}$$

Bewijs

Het is eenvoudig in te zien dat de $T > t$ a.s.a. er nul voorkomens zijn in het interval $[0, t]$:

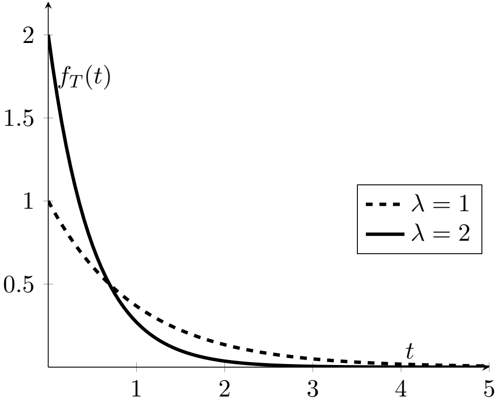
$$\mathbb{P}(T > t) = \mathbb{P}(\text{nul voorkomens in}[0, t]) = e^{-\lambda t}.$$

Uit de complementswet volgt dan dat

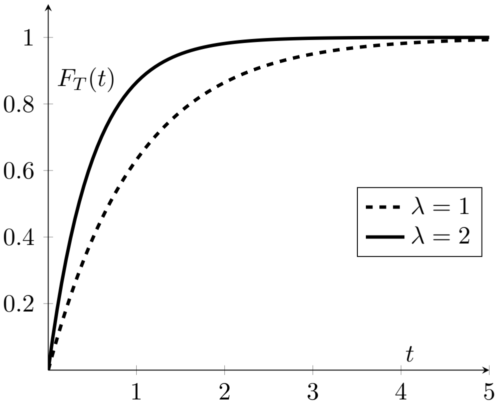
$$F_T(t) = \mathbb{P}(T \leq t) = 1 - e^{-\lambda t}.$$

Door het berekenen van de afgeleide functie (naar t) vindt men onmiddellijk de formule voor $f_T(t)$.

Kansdichtheid: grafisch



Kansverdelingsfunctie: grafisch



Verwachtingswaarde en variantie

Eigenschap

Als T een toevalsveranderlijke is die exponentieel verdeeld is met parameter λ dan geldt

$$\mu_T = \frac{1}{\lambda}$$

en

$$\sigma_T^2 = \frac{1}{\lambda^2}.$$

Opmerkingen:

- Hoe groter de dichtheid van de voorkomens (i.e. hoe groter λ) hoe kleiner de gemiddelde wachttijd.
- Het resultaat is zeer logisch: als er bv. (gemiddeld) 5 gebeurtenissen per minuut zijn (i.e. $\lambda = 5/\text{minuut}$), dan duurt het (gemiddeld) $1/5$ minuut voor een gebeurtenis zich voordoet.

Markov-eigenschap

Eigenschap

De exponentiële verdeling voldoet aan de MARKOV *eigenschap; ze bezit namelijk geen geheugen. Wiskundig betekent dit dat voor elke grootheid* Δt *geldt dat*

$$\mathbb{P}(T > t + \Delta t \,|\, T > \Delta t) = \mathbb{P}(T > t).$$

Opmerking:

Het is dus alsof de exponentiële verdeling "vergeet" hoelang er reeds gewacht is.

Bewijs

Het bewijs is een rechtstreekse toepassing van de formule voor voorwaardelijke kans:

$$
\begin{aligned}
\mathbb{P}(T > t + \Delta t \,|\, T > \Delta t) &= \frac{\mathbb{P}\big((T > t + \Delta t) \cap (T > \Delta t)\big)}{\mathbb{P}(T > \Delta t)} \\
&= \frac{\mathbb{P}(T > t + \Delta t)}{\mathbb{P}(T > \Delta t)} \\
&= \frac{e^{-\lambda(t+\Delta t)}}{e^{-\lambda \Delta t}} \\
&= e^{-\lambda t} \\
&= \mathbb{P}(T > t).
\end{aligned}
$$

Oefeningen

1. Jobs worden naar een printer gezonden aan een gemiddeld tempo van 3 jobs per uur.
   1.1 Wat is de verwachte tijd tussen jobs?
   1.2 Wat is de waarschijnlijkheid dat de volgende job wordt gezonden binnen de volgende 5 minuten?
2. De tijd nodig door een technicus om een bepaald soort machine te herstellen heeft een exponentiële verdeling met een gemiddelde van 4 uur. Er bestaat speciaal gereedschap dat dit gemiddelde doet dalen naar 2 uur. Wanneer de technicus erin slaagt om een machine te herstellen in minder dan 2 uur, dan krijgt hij 100 EUR, anders krijgt hij 80 EUR. Bepaal de verwachte meeropbrengst per machine wanneer de technicus het speciale gereedschap gebruikt.

Oefeningen

3. Een programma is verdeeld in 3 modules die parallel worden gecompileerd op 3 verschillende computers. De compilatietijd van elk blok is exponentieel verdeeld met een gemiddelde van 5 minuten, onafhankelijk van de andere blokken. Het programma is gecompileerd wanneer alle blokken gecompileerd zijn.
   3.1 Wat is de waarschijnlijkheid dat het programma volledig gecompileerd is in minder dan 5 minuten?
   3.2 Wat is de verwachte compilatietijd? Maak gebruik van het feit dat je de verwachtingswaarde van een toevalsveranderlijke X die enkel niet-negatieve waarden aanneemt ook als volgt kan berekenen:
   $$\mu_X = \int_0^\infty (1 - F_X(x))\,\mathrm{d}x.$$